

10/525696

PCT/PTO 22 FEB 2005

TH

REC'D 20 OCT 2003

WIPG BET

P1 1079368

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

October 15, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/408,379

FILING DATE: September 05, 2002

RELATED PCT APPLICATION NUMBER: PCT/US03/28076



By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

M. Sias

M. SIAS
Certifying Officer

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

BEST AVAILABLE COPY

09/05/02
Jc931 U.S. PTO

09-06-02

Please type a plus sign (+) inside this box → ☒

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

I hereby certify that this paper is being deposited with the United States Postal Service as Express Mail in an envelope addressed to: Box Provisional Application, Asst. Comm. for Patents, Washington, D.C. 20231, on this date

SEP 5 2002

U.S. Patent or Date

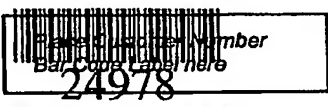
Express Mail Label No.: EL846176456US

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EL846176456US

Alper

INVENTOR(S)					
Given Name (first and middle [if any])		Family Name or Surname		Residence (City and either State or Foreign Country)	
Rene L. Arvind V.		Cruz Santhanam		LaJolla, CA	
<input type="checkbox"/> Additional inventors are being named on the ___ separately numbered sheets attached hereto					
TITLE OF THE INVENTION (280 characters max)					
OPTIMAL LINK SCHEDULING AND POWER CONTROL FOR CDMA WIRELESS MULTI-HOP NETWORKS					
Direct all correspondence to: CORRESPONDENCE ADDRESS					
<input checked="" type="checkbox"/> Customer Number		24978			
OR		Type Customer Number here			
<input type="checkbox"/> Firm or Individual Name		PATENT TRADEMARK OFFICE			
Address					
Address					
City		State	ZIP		
Country		Telephone	Fax		
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/> Specification Number of Pages		17		<input type="checkbox"/> CD(s), Number	
<input type="checkbox"/> Drawing(s) Number of Sheets				<input type="checkbox"/> Other (specify)	
<input type="checkbox"/> Application Data Sheet. See 37 CFR 1.76					
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT					
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.				FILING FEE AMOUNT (\$)	
<input checked="" type="checkbox"/> A check or money order is enclosed to cover the filing fees				80.00	
<input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number:		07-2069			
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.					
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input type="checkbox"/> No.					
<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____					

Respectfully submitted,

SIGNATURE

TYPED or PRINTED NAME Steven P. Fallon

TELEPHONE (312) 360-0080

Date 9/5/02

REGISTRATION NO.
(if appropriate)
Docket Number:

35,132

0321.66895

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the complete provisional application to the PTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, Washington, D.C. 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C. 20231.

Sep. 5, 2002
Date

Dai D. O'Connell
Express Mail Label No.: EL846176456US

OPTIMAL LINK SCHEDULING AND POWER CONTROL FOR CDMA WIRELESS MULTI-HOP NETWORKS

Consider a generic multi-hop network with a set of wireless devices (abstracted as nodes) communicating with each other in a certain pre-specified configuration (abstracted as directed links). Nodes could represent laptops, PDA's, Access Points etc., while directed links indicate data transfer between nodes. All nodes share a common frequency spectrum and use DS-CDMA, transmitting their data using a signature sequence. This results in each link's transmission interfering with other ongoing transmissions in the network. We have developed an algorithm that maximizes the weighted sum of long-term average link rates (throughput) of a multi-hop network when each link in the network is subject to peak and average power constraints. We call any subset of the links in the given configuration a transmission mode. Given the set of directed links and the signal attenuation between transmitters and receivers in the network, our proposed algorithm determines the throughput maximizing strategy. Specifically, it computes the optimal set of transmission modes, the relative duration for which each transmission mode is active and the power of signal transmission of every transmitter in each transmission mode. For more details refer to the attached manuscript.

Using essentially the same methodology, an algorithm has been developed that minimizes a weighted sum of the average power used by transmitters, subject to a lower bound on the average data rate on each link, and peak power constraints per transmitter. The algorithm yields a schedule and power allocations for each transmitter. Moreover, an additional output of the algorithm is a sensitivity value, for each link, which specifies the marginal increase in average (weighted) power required for a marginal increase in the average data rate on the given link. These link sensitivities can be used as link weights in a shortest path routing algorithm, enabling energy efficient routing. Although the methodology was conceived for DS-CDMA networks, the same methodology could be applied in wireless FDMA networks as well.

Existing approaches pre-specify the link scheduling policy based on a graph-theoretic approach. They schedule links that are geographically far away (virtually interference-free) from each other simultaneously. e.g. maximum independent set of nodes in a graph. In our approach, we explicitly account for multi-access interference and determine the throughput maximizing transmission strategy. In fact we find that as the level of ambient noise increases, the optimal policy schedules links in close spatial proximity simultaneously. A distinguishing feature of our approach is in the use of convex duality to solve this resource allocation problem. We establish that optimizing transmissions over the infinite time horizon can be achieved by optimizing over every single slot. Next, we compute the optimal subsets of links by solving the reduced dual problem over a single slot. We have developed an iterative descent algorithm that solves the dual problem in a finite number of steps. Solving the dual problem yields us a small

set of transmission modes. Finally, we determine the optimal time-sharing of these transmission modes to maximize network capacity.

Given a set of wireless devices and a certain pre-specified configuration of links, sharing a common frequency spectrum, our algorithm specifies a transmission strategy maximizes the weighted sum of long-term link rates when each link is subject to peak and average power constraints. The main input to our algorithm is the signal attenuation between all transmitters and receivers in the network.

We solve the above constrained optimization problem using the theory of convex duality. First, we establish that finding the best transmission strategy over the infinite time horizon can be achieved by optimizing over every single slot. Next, we compute the optimal transmission modes by solving the reduced dual problem over a single slot. We have developed an iterative descent algorithm that solves the reduced dual problem in a finite number of steps. Solving the dual problem yields us a small set of optimal transmission modes. Finally, we determine the optimal time-sharing of these transmission modes to maximize network capacity through a series of matrix inversions, each matrix sparse and comprising of either 1's or 0's.

Optimal Link Scheduling and Power Control in CDMA Multihop Wireless Networks

Rene L. Cruz & Arvind V. Santhanam

Abstract—

In this paper, we study the problem of radio resource allocation in wireless multihop networks. The main objective is to develop integrated link scheduling and power control policies to maximize the network throughput. Transmitters in the network are subject to peak and average power constraints. Subject to these constraints, we find optimal subsets of concurrent transmitters and associated power levels to maximize network throughput, taking into account multi-access interference. We find that when ambient noise levels are sufficiently high, it is generally desirable to enable simultaneous transmissions, even if they are in close geographic proximity. We compare our results against pure CDMA to illustrate the significant gains of our approach. The results in this paper are applicable to multihop and single hop networks.

Index Terms—CDMA, Multihop Networks, Resource Allocation.

I. INTRODUCTION

In recent years, there has been a great deal of interest and development in supporting data applications in multi-access wireless networks. Wireless devices are however, severely limited in bandwidth and energy and must use their resources judiciously. Resource allocation policies in currently deployed systems like IEEE802.11 though simple and distributed in nature, perform poorly even under moderate loads. Centralized scheduling of node transmissions can greatly improve network capacity, e.g see [3] [4]. Consistent with this, the IEEE802.11 standard includes a *point coordination function* (PCF) which centrally schedules users. The High Data Rate (HDR) standard also schedules transmissions centrally on the down-link channel. The benefits of centralized scheduling for multihop networks are equally rewarding if not more. In this paper, we extend link scheduling from one-hop networks to the multihop paradigm.

The capacity of a wireless channel is a function of the signal to interference plus noise ratio (SIR) of that channel. For instance, the Shannon capacity of a link i over a frequency bandwidth W is $C(i) = W \log_2(1 + \text{SIR}(i))$

(assuming the noise is Gaussian). The dependence of link rates on powers naturally raises the possibility of improving capacity through a suitable choice of powers. Before delving into specifics, we define some terms commonly used in this paper. A directed link $l \in \{i, j\}$ in a multihop network of N nodes facilitates data transfer from node i to node j . We call any subset of L_a potential links in a network a *transmission mode*. There are 2^{L_a} transmission modes for an L_a link network. A list of transmission modes and the relative time-fractions for which they are “used” comprises a link scheduling policy. For each transmission mode, we also need to specify the power used by active transmitters, which constitutes a power control policy.

The conventional approach to scheduling links in a multihop networks is to allow nodes far away from each other to transmit simultaneously. Such a link scheduling policy ensures essentially interference-free transmissions, thereby achieving high instantaneous individual link rates. Nodes in the *vicinity*¹ of each other have an edge in the channel contention graph G_{ch} . Computing a maximum independent set [1] in the channel contention graph requires global topology information, entailing centralized processing. Good approximation algorithms that allow for distributed implementation would naturally be the best recourse. One such example is the *minimum-degree greedy algorithm* [2], a heuristic that chooses nodes with minimum degrees in subgraphs.

In our approach, we explicitly model the effects of signal interference, allowing concurrent transmissions in a close geographic proximity. By globally optimizing the transmission modes and the associated power levels, we can achieve much higher throughputs than previous approaches. Specifically, we solve the following resource allocation problem:

Problem: Find an optimal link scheduling and power control policy to maximize the network throughput, when transmitters are subject to peak and average power constraints.

The main input to our algorithm is the path-loss matrix of the entire network. This matrix consists of N^2

The authors are with the Dept. of Electrical and Computer Engg., UCSD, USA. E-mail: asanthan@ucsd.edu.

¹Typically within a 2-hop distance of each other.

gain values, one between each pair of nodes. In solving this problem, we only specify the relative time-fractions for which each transmission mode is active, not the exact time when it is "on". Viewing the time-fractions as weights, transmission modes can be scheduled over slots, using algorithms like deficit round robin.

II. SYSTEM MODEL

Consider a system with N nodes in a CDMA environment. A subset of L_a transmission links (among the possible $N(N-1)$ links) constitutes a network topology, and is assumed to be prespecified. We wish to maximize the weighted sum of throughputs on these links, as will be described below in equation (3). For a given link $l = \{i, j\}$, the transmitter node i uses a signal power $P(l)$. The path gain from node i to node j is given by $G(i, j)$. The path gain $G(i, j)$ models the effects of signal attenuation due to distance, channel fading and shadowing. We assume that the path gains $G(i, j)$ are constant. Denoting the transmitting and receiving nodes of link l by $T(l)$ and $R(l)$ respectively, the signal to interference and noise ratio (SIR or $\gamma(l)$) for link l is given by,

$$\gamma(l) = \frac{G(T(l), R(l))P(l)}{\sum_{k:k \neq l} P(k)G(T(k), R(l)) + n_{R(l)}} \quad (1)$$

where, $n_{R(l)}$ is the ambient noise power at node $R(l)$. Let the raw data rate of link l be $X(l)$. We assume $X(l)$ is a linear function of its SIR and the total available bandwidth W . Assuming the Gaussian approximation [6] to compute the bit-error-rate (BER), the data rate of a link with a tolerable BER of 10^{-q} and using BPSK modulation is given by

$$X(l) = W' \gamma(l) \quad \text{where} \quad W' = \frac{W}{2q \log_2 10}. \quad (2)$$

In fact, we consider a time slotted model, where all slots are of equal duration and transmissions begin and end on slot boundaries. We let $X_m(l)$ and $P_m(l)$ be the data rate for link l and transmission power for $T(l)$ in slot m . We then have

$$X_m(l) = W' \left(\frac{G(T(l), R(l))P_m(l)}{\sum_{k:k \neq l} G(T(k), R(l))P_m(k) + n_{R(l)}} \right).$$

The long-term average rate of link l is then defined as

$$X^{avg}(l) = \liminf_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m X_k(l).$$

The average power consumed by each link l , $P^{avg}(l)$ is the time-average over an infinite horizon :

$$P^{avg}(l) = \limsup_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m P_k(l).$$

Let the weights associated with each link be $\alpha(l)$. These weights are a reflection of the relative priority of links. We wish to maximize the average throughput (capacity) of the network, $f(\vec{X}^{avg})$, which is simply the weighted sum of long-term average rates:

$$f(\vec{X}^{avg}) = \sum_{l=1}^{L_a} \alpha(l) X^{avg}(l). \quad (3)$$

The transmitter for each link l is constrained to expend an average power of at most $\bar{P}^{avg}(l)$ and a peak power of at most $P^{max}(l)$.

A. Primal Problem

The objective of this paper is to compute transmission schedules for the links in the system so as to maximize the system throughput of the wireless network when links are subject to peak and average power constraints. The primal problem can be succinctly stated as,

$$\{\max f(\vec{X}^{avg}) \text{ subject to } P^{avg}(l) \leq \bar{P}^{avg}(l) \text{ and } 0 \leq P_m(l) \leq P^{max}(l)\} \quad (4)$$

Note this optimization involves choosing optimal power levels for each transmitter in each slot. Furthermore, the achieved instantaneous throughput in each slot is not a convex function of the transmission powers in the slot. However, using a duality approach below, we reduce the problem to a convex optimization problem over a single slot.

B. Duality Approach

We exploit a *convex duality* approach to solve the primal problem. Defining a set of dual variables $\vec{\beta} = \{\beta(1), \beta(2) \dots \beta(L_a)\}$ for each link, the dual objective function is

$$g(\vec{\beta}) = \max \{ f(\vec{X}^{avg}) + \sum_{l=1}^{L_a} \beta(l) (\bar{P}^{avg}(l) - P^{avg}(l)) \} \\ \text{subject to } 0 \leq \bar{P}_m \leq \bar{P}^{max} \quad (5)$$

Note the absence of the average power constraints in the definition of $g(\vec{\beta})$. For any non-negative vector $\vec{\beta}$, it is

clear that the maximum value of the objective function in the optimization problem (4) is upper bounded by $g(\vec{\beta})$. This observation leads to the dual optimization problem,

$$\min g(\vec{\beta}) \quad \text{subject to} \quad \{\vec{\beta} \geq 0\} \quad (6)$$

Computation of $g(\vec{\beta})$ involves optimizing over the set of powers $\{0 \leq \vec{P}_k \leq \vec{P}^{\max}\}$ in every single slot k . Using the linearity of $f(\cdot)$ and the fact that the definition of $g(\vec{\beta})$ does not directly incorporate average power constraints, it is clear that $g(\vec{\beta})$ can be evaluated by optimizing over a single slot. i.e. $g(\vec{\beta}) = h(\vec{\beta})$ where we define $h(\vec{\beta})$ below. Defining

$$Q(\vec{P}_1, \vec{\beta}) = \sum_{l=1}^{L_a} [\alpha(l)X_1(l) + \beta(l)(\bar{P}^{avg}(l) - P_1(l))] \quad (7)$$

the dual function over a single slot $h(\vec{\beta})$ is simply

$$h(\vec{\beta}) = \max_{0 \leq \vec{P}_1 \leq \vec{P}^{\max}} Q(\vec{P}_1, \vec{\beta})$$

Therefore we consider the simplified dual problem

$$\min h(\vec{\beta}) \quad \text{subject to} \quad \{\vec{\beta} \geq 0\} \quad (8)$$

III. SOLVING THE DUAL PROBLEM

In order to solve the dual problem, we must first compute the dual objective function. Consider the polyhedral set $S_P = \{\vec{P}_1 : 0 \leq \vec{P}_1(l) \leq P^{\max}(l), l = 1, 2, \dots, L_a\}$. Let $\vec{P}_m^{\text{ext}}, m = 1, 2, \dots, 2^{L_a}$ be the extreme points of S_P , i.e. each point in S_P can be uniquely represented as a convex combination of $\vec{P}_m^{\text{ext}}, m = 1, 2, \dots, 2^{L_a}$. Using the fact that $Q(\vec{P}_1, \vec{\beta})$ for a fixed $\vec{\beta}$, is convex in each variable $P_1(l)$, the following lemma can be proven

Lemma 1:

$$h(\vec{\beta}) = \max\{Q(\vec{P}_m^{\text{ext}}, \vec{\beta}) : m = 1, 2, \dots, 2^{L_a}\} \quad (9)$$

It is interesting to note that $Q(\vec{P}_1, \vec{\beta})$ is not a convex function of \vec{P}_1 , only convex in each coordinate of \vec{P}_1 .

A. Minimizing the Dual Objective function

The dual objective function $g(\vec{\beta})$ is the *pointwise maximum* of 2^{L_a} affine functions. Therefore, $h(\vec{\beta})$ is a piecewise linear (convex) continuous function of $\vec{\beta}$.

$$h(\vec{\beta}) = \max_{1 \leq m \leq 2^{L_a}} \{a_m \vec{\beta} + b_m\} \quad (10)$$

where $a_m = \bar{P}^{avg} - \vec{P}_m^{\text{ext}}$ and $b_m = Q(\vec{P}_m^{\text{ext}}, \vec{0})$. We have developed an iterative descent algorithm to solve the dual problem (8). Exploiting the piecewise linearity of $h(\vec{\beta})$, our algorithm converges in a finite number of iterations. Alternately, by introducing a slack variable, (8) can be reduced to a linear program and be efficiently solved using interior point methods. Solving the dual problem yields $\vec{\beta}^*$ and $P^{*,i}$, such that

$$\min\{h(\vec{\beta}) : \vec{\beta} \geq 0\} = h(\vec{\beta}^*) = Q(\vec{P}^{*,i}, \vec{\beta}^*) \quad \text{for } i \in [1, \hat{K}]$$

The vectors $\vec{P}^{*,i}$ are extreme points of S_P and are called the optimal transmission modes.

IV. OPTIMAL LINK SCHEDULING POLICY

The solution to the dual problem yields us a set of optimal transmission modes. The complementary slackness condition ensures that links with $\beta^*(l) > 0$ meet their average power constraints with equality. At least one user will meet his average power constraint with equality in the optimal solution. Suppose L_{eq} is defined as $L_{eq} = \{l : \beta^*(l) > 0\}$, and let $M = |L_{eq}|$. Using the principle of strong duality [5], there exists $\lambda^* = \{\lambda^*(1), \lambda^*(2), \lambda^*(\hat{K})\}$ such that the equations below are satisfied.

$$\begin{aligned} \sum_{i=1}^{\hat{K}} \lambda^*(i) P^{*,i}(l) &= P^{avg}(l) \quad \text{for all links } l \in L_{eq} \\ \sum_{i=1}^{\hat{K}} \lambda^*(i) P^{*,i}(l) &< P^{avg}(l) \quad \text{for all links } l \notin L_{eq} \\ \sum_{i=1}^{\hat{K}} \lambda^*(i) &= 1 \quad \text{where } \lambda^*(i) \geq 0 \end{aligned}$$

Note that \hat{K} is not necessarily equal to $M + 1$, so the above system of equations may be under or over determined. If $\hat{K} \leq M + 1$, we compute λ^* by inverting a $\hat{K} \times \hat{K}$ matrix. If $\hat{K} > M + 1$, we invert a series of $R_{max} \times R_{max}$ matrices, corresponding to different subsets of optimal transmission modes, where the cardinality of the subsets are equal to $R_{max} \leq M + 1$. Note that the matrices involved have elements which are either 0 or 1, which simplifies the computation. We terminate upon finding a solution, which is guaranteed by the principle of strong duality.

The optimal policy consists of time sharing between the optimal transmission modes, where the fraction of time that transmission mode $\vec{P}^{*,i}$ is used is equal to $\lambda^*(i)$ for each $i \in [1, \hat{K}]$. In every slot, if a transmitter is included in the enabled transmission mode, it transmits at the peak power level allowed for that transmitter.

V. SIMULATION RESULTS AND DISCUSSION

In this section we compare the performance of our algorithm with CDMA over the uplink channel for a single cell (i.e. single hop) network of 8 nodes. Nodes are scattered uniformly in a circular region of radius 200 meters. The peak power of links are $P^{\max} = 1$ Watt. Links have equal weights i.e. $\alpha(l) = \frac{1}{8}$. We compare the throughput achieved by the optimal policy against a CDMA policy as a function of ambient noise (normalized to the peak received power) and the average to peak power ratio. All users in the CDMA policy transmit at their average power \bar{P}_{avg} every slot. Figure 1 depicts the ratio of the throughput of our policy to that of the CDMA policy as a function of ambient noise for different values of average to peak power ratio, namely $\bar{P}_{\text{avg}} = \frac{P^{\max}}{7}$, $\frac{P^{\max}}{5}$ and $\frac{P^{\max}}{2}$.

In low noise regimes, the optimal policy outperforms the CDMA policy by at least a factor of 300 for all average to peak power ratios considered. The optimal policy in low noise regimes schedules good links one at a time in TDMA sequence. In this regime, CDMA performs poorly due to high mutual interference between transmissions. As the ambient noise level increases, the optimal policy schedules multiple simultaneous transmissions for unequal amounts of time. In fact in high noise regimes, the optimal policy schedules a large number of simultaneous transmissions; the throughput achieved by the CDMA policy therefore approaches that of the optimal policy.

In low noise regimes, we observe a high improvement in throughput with increasing average to peak power ratios. This is so because the optimal policy schedules good links to transmit with increasing frequency (almost equal to their peak to average power ratios). At higher ambient noise levels, there is relatively little improvement in throughput for higher average to peak power ratios. This is so because the optimal policy schedules a large number of simultaneous transmissions. The optimal policy uses lower average power than the specified limit in all noise regimes. In fact, in low noise regimes, the average power consumed by the optimal policy is significantly less.

VI. CONCLUSION

We developed an integrated link scheduling and power allocation policy for a general multihop network that maximizes network throughput subject to peak and average power constraints, assuming constant channels. We believe we can extend our algorithms to the case of block fading channels. The power control policy is extremal and therefore simple. The capacity is achieved by time-sharing a small number of transmission modes ($\leq L_a + 1$).

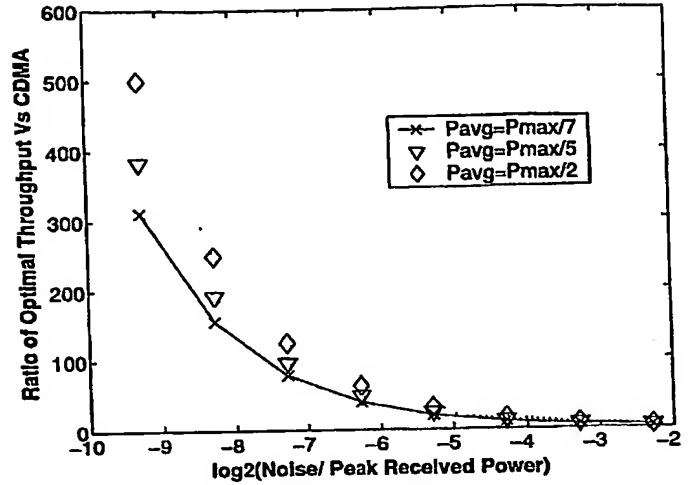


Fig. 1. Optimal Policy vs CDMA for low values of ambient noise.

The set of optimal transmission modes depend on the level of ambient noise and average-to-peak power ratios. The optimal policy in low noise regimes is to schedule links with good channels and low average to peak power ratios, in a TDMA sequence. In moderate or high noise regimes, the optimal strategy involves scheduling multiple simultaneous transmissions even though they may be in close geographic proximity.

REFERENCES

- [1] A. Ephremides, T. V. Truong, "Scheduling Broadcasts in multihop radio networks", *IEEE Trans. on Communications* v.38, Apr. 1990.
- [2] H. Luo, S. Lu, V. Bharghavan, "A New Model for Packet Scheduling in Multihop Wireless Networks", *ACM Mobicom*, August 2000.
- [3] S. Toumpis, A. J. Goldsmith, "Capacity Regions for Wireless Adhoc Networks", *International Symposium on Communication Theory and Applications*, April 2001.
- [4] R. Knopp, P. A. Humblet, "Information capacity and power control in single-cell multiuser communications", *International Conf. on Communications*, June 1995.
- [5] D. Bertsekas, "Non-linear Programming", *Athena Scientific*, 1999.
- [6] T. S. Rappaport, "Wireless Communications: Principles and Practice", *Prentice Hall*.

Optimal Routing, Link Scheduling and Power Control in Multi-hop Wireless Networks

R. L. Cruz and Arvind V. Santhanam
 Department of Electrical & Computer Engineering
 University of California, San Diego
 La Jolla, CA 92093-0407
 cruz@ece.ucsd.edu , arvind@cw.cw.ucsd.edu

Abstract—

In this paper, we study the problem of joint routing, link scheduling and power control to support high data rates for broadband wireless multi-hop networks. We first address the problem of finding an optimal link scheduling and power control policy that minimizes the total average transmission power in the wireless multi-hop network, subject to given constraints regarding the minimum average data rate per link, as well as peak transmission power constraints per node. Multi-access signal interference is explicitly modeled. We use a duality approach whereby, as a byproduct of finding the optimal policy, we find the sensitivity of the minimal total average power with respect to the average data rate for each link. Since the minimal total average power is a convex function of the required minimum average data rates, shortest path algorithms with the link weights set to the link sensitivities can be used to guide the search for a globally optimum routing. We present a few simple examples that show our algorithm can find policies that support data rates that are not possible with conventional approaches. Moreover, we find that optimum allocations do not necessarily route traffic over minimum energy paths.

Index Terms—Wireless Networks, Multi-hop Networks, Routing, Scheduling, Power Control.

I. INTRODUCTION AND MODEL

Broadband wireless networks today are capable of supporting high data rates. Minimizing the total power in such systems is of paramount importance not just to increase its own operational lifetime in the case of battery powered devices, but also to coexist symbiotically with other systems which share the same frequency spectrum. For instance, 802.11 LANs and Bluetooth networks share the same unlicensed band and can mutually benefit by limiting the power of their respective signal transmissions.

Most currently deployed networks that support high data rates use only a “single hop” for wireless communication between a mobile user and an access point that is connected to the wired infrastructure. However, the use of multiple hops to transport data has been shown to enhance

network capacity [9] [8] [15], and may be necessary due to cabling limitations in many environments.

Consider a network of base-stations interconnected to each other through wireless links. Each base-station serves as an ingress or egress for the aggregate of traffic associated with the mobile users in its domain, and routes its data through other base stations via multiple wireless hops to an access point connected to a wired infrastructure. In this paper, we focus on the transport of such aggregated traffic between base stations. We present algorithms to compute the optimal routing, link scheduling and power control policy that consumes minimum total average power. While current approaches tend to decouple routing, link scheduling and power control from each other, we take an integrated approach.

Suppose there are N stationary nodes, labeled by the integers $1, 2, \dots, N$, which correspond to the base stations discussed in the previous paragraph. A set \mathcal{E} of $L_{\mathcal{E}} = |\mathcal{E}|$ transmission links, among the possible $N(N - 1)$ links between nodes, constitutes a network topology. For now, assume that the transmissions on all of these links occur concurrently.

For a given link $l = (i, j)$, the transmitter node i uses a signal power $P(l)$. The path “gain” from node i to node j is given by $G(i, j)$, and models the effects of signal attenuation due to distance, channel fading and shadowing, as well as antenna gain patterns. We assume that the path gains $G(i, j)$ are constant. The transmitting and receiving nodes of link l are denoted by $T(l)$ and $R(l)$ respectively. The received signal power at node $R(l)$ from the transmitter $T(l)$ thus given by $P(l)G(T(l), R(l))$. However, signals emanating from other transmitters appear to the receiver $R(l)$ as interference, and there is ambient thermal noise to contend with as well. The signal to interference and noise ratio (SIR) for link l is defined as

$$\gamma(l) = \frac{G(T(l), R(l))P(l)}{\sum_{k: k \neq l} P(k)G(T(k), R(l)) + n_{R(l)}}, \quad (1)$$

where n_j is the ambient noise power at node j .

The capacity of the single wireless channel embodied by link l can be modeled as a function of the SIR, $\gamma(l)$. For instance, the Shannon capacity of a link l over a frequency bandwidth W is $W \log_2(1 + \gamma(l))$, assuming Gaussian noise and interference. Throughout this paper, we assume that the achieved data rate on link l , $X(l)$, is a linear function of the SIR, $\gamma(l)$, and the total available bandwidth W . This is a reasonable approximation to the Shannon capacity at small values of SIR, which for example is the case in CDMA networks where all links share a common frequency band and have at least a moderate processing gain. Our model could easily encompass the case where the links may use different frequency bands, and in this case the path gains incorporate crosstalk due to imperfect bandpass filters. However, for concreteness throughout this paper, we shall assume all links share the same frequency band of width W . Assuming the Gaussian approximation [12] to compute the bit-error-rate (BER), the data rate of a link l with a tolerable BER of 10^{-9} and using BPSK modulation is given by

$$X(l) = W' \gamma(l) \quad \text{where} \quad W' = \frac{W}{2q \log_2 10}. \quad (2)$$

Suppose that the desired data rate on each link l is $C(l)$ bits per second, i.e.

$$X(l) \geq C(l), \quad l \in L_E. \quad (3)$$

Using (1) and (2), we can rewrite (3) in matrix form as

$$\mathbf{P} \geq \mathbf{F}\mathbf{P} + \mathbf{b}, \quad (4)$$

where $\mathbf{P} = (P(1), \dots, P(L_E))^T$,

$$b(l) = \frac{C(l)n_{R(l)}}{G(R(l), T(l))W'},$$

$\mathbf{b} = (b(1), \dots, b(L_E))^T$, and \mathbf{F} is the $L_E \times L_E$ matrix with entries

$$F(l, k) = \begin{cases} \frac{G(R(l), T(k))C(l)}{G(R(l), T(l))W'}, & \text{if } k \neq l \\ 0, & \text{if } k = l. \end{cases}$$

If \mathbf{F} is a stable matrix, i.e. all eigenvalues of \mathbf{F} are strictly inside the unit circle, then $(\mathbf{I} - \mathbf{F})^{-1}$ exists and has non-negative elements. Thus, in this case, we can rearrange (4) and premultiply the resulting matrix inequality by $(\mathbf{I} - \mathbf{F})^{-1}$ to obtain

$$\mathbf{P} \geq (\mathbf{I} - \mathbf{F})^{-1} \mathbf{b} \quad (5)$$

Hence, we can view $(\mathbf{I} - \mathbf{F})^{-1} \mathbf{b}$ as the minimal power vector that supports the network topology defined by links \mathcal{E} and their associated required data rates. Foschini and Miljanic [7] proposed a simple distributed algorithm to compute the minimum power vector in a cellular network context. Subsequently, Bambos, Chen, and Pottie [1] considered a distributed link admission control algorithm for the case of multi-hop networks, in the framework of [7].

If the eigenvalues of \mathbf{F} are on or outside the unit circle, there does not exist any power vector \mathbf{P} which supports the required data rates. If the eigenvalues of \mathbf{F} are inside the unit circle but close to the boundary, the minimal power vector $(\mathbf{I} - \mathbf{F})^{-1} \mathbf{b}$ will be very large. Even if the eigenvalues of \mathbf{F} are well inside the unit circle, using the minimal power vector indicated in (5) may be inefficient. This is illustrated in the following example.

Example: Consider a square network of 4 nodes and 2 links $\{(1, 2), (3, 4)\}$. Suppose the relevant path gain values are $G(1, 2) = G(3, 4) = G_0$ and $G(1, 4) = G(3, 2) = G_0/2$, corresponding to omnidirectional antennas and inverse square law propagation loss. Let the required data rates be $X(\{1, 2\}) = X(\{3, 4\}) = \frac{W'}{2}$, and let the ambient noise n_0 be the same at both the receivers, specifically $n_0 = G_0$. In this case, it is readily verified that $(\mathbf{I} - \mathbf{F})^{-1} \mathbf{b} = (0.67, 0.67)^T$, i.e. each transmitter is required to continuously transmit at a power of 0.67 Watt. On the other hand, if instead the transmitters take turns transmitting at a power of 1 Watt, such that only one transmitter is active at any time, the peak data rate that can be achieved is twice that of the data rate achieved previously, due to a lack of interference. Thus, if the transmitters are active for 50% of the time, the same average data rate of $\frac{W'}{2}$ is achieved, but this consumes an average power of only 0.5 Watt, which is 33% less than before.

This example suggests that it can be inefficient to operate with all links in a wireless network active concurrently. Indeed, in most wireless networks, links operate in a half-duplex mode to limit interference. We can essentially eliminate interference by scheduling only links which are very far apart, as is proposed for example in [6][10]. However, in order to achieve high data rates, this may require unacceptably high peak transmission power. Constraints on peak transmission power may necessitate the use of concurrent transmissions in this case.

Recently, Elbatt and Ephremides [5] proposed a scheduling and power control strategy that selects candidate subsets of concurrently active links, and applies the power control algorithm in [7] in order to find the mini-

mal vector $(\mathbf{I} - \mathbf{F})^{-1}\mathbf{b}$ for each candidate subset. The objective of their optimization problem is to maximize network throughput by supporting as many links as possible with a given fixed SIR, while remaining energy efficient. They propose a heuristic to identify large subsets of concurrently active links that can be efficiently supported, and evaluate the throughput of the heuristic via simulation for a uniform traffic demand on links.

In this paper, we propose a systematic optimization for a problem formulation that is motivated by similar objectives. In particular, we seek to find subsets of simultaneously active links, herein called *transmission modes*, as well as the associated transmission powers, in order to minimize the total *average* transmission power expended across the network. The constraints of this optimization problem are that each link in the network must have an average data rate no less than a given prescribed value, and a peak transmission power for each node. This problem statement is formalized in Section II, and a duality approach for finding the optimal scheduling and power control policy is presented.

The required minimum average data rates on each link are ultimately determined by the average rates at which traffic is generated by users and the routing algorithm that is used. In wireless networks it is possible to reconfigure the data rates of the links on a fast time scale in response to changing traffic and channel conditions. In this paper, we shall also consider the problem of routing, and hence determination of the required data rates on each link, for a given traffic demand rate matrix. We shall address this in Section III, exploiting the duality approach explained in Section II.

After presenting the details of our approach, we discuss several illustrative examples in Section IV. Qualitatively, we find that in high ambient noise regimes, a large number of links in close geographic proximity are active simultaneously in optimal policies. In addition, we find that route diversity allows nodes to split traffic over multiple paths and support higher throughputs than could otherwise be achieved. This is somewhat surprising since all links share a common bandwidth. However, pushing the data rates to the maximum possible limit comes at price: As the data rates increase, the network tends to become less energy efficient.

We conclude the paper with some closing remarks in Section V.

II. OPTIMAL SCHEDULING AND POWER CONTROL

In this section, we formally specify the optimization problem discussed in the previous section, and present a

duality approach for solving it. We then discuss an approach to limit computational complexity.

For simplicity of exposition, we divide time into slots, each of equal duration and indexed by the positive integers. Transmissions begin and end on slot boundaries. Generalizing the notation introduced earlier, let $X_m(l)$ and $P_m(l)$ be the data rate for link l in slot m , and transmission power for the transmitter $T(l)$ for link l in slot m , respectively. Let $\vec{P}_m = (P_m(1), P_m(2), \dots, P_m(L_E))$ be the network power vector for slot m . Let $P^{max}(i)$ be the maximum transmission power for node i . Let $\mathcal{E}(i)$ be the links in \mathcal{E} that originate at node i .

Each node must conform to the peak transmission power constraint in every slot:

$$\begin{aligned} 0 &\leq \sum_{l \in \mathcal{E}(i)} P_m(l) \leq P^{max}(i) \text{ and} \\ 0 &\leq P_m(l), \text{ for all } m \geq 1. \end{aligned} \quad (6)$$

Using (2) and (1), the achieved data rate for link l in slot m is

$$X_m(l) = W' \left(\frac{G(T(l), R(l)) P_m(l)}{\sum_{k: k \neq l} G(T(k), R(l)) P_m(k) + n_{R(l)}} \right).$$

The long-term average rate of link l is then defined as

$$X^{avg}(l) = \liminf_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m X_k(l).$$

For each link l , let $C(l)$ be a given minimum acceptable average data rate, i.e. we must have

$$X^{avg}(l) \geq C(l), \text{ for all } l \in \mathcal{E}. \quad (7)$$

Define the required minimum average rate vector as $\vec{C} = (C(1), C(2), \dots, C(L_E))$.

The average power consumed by the transmitter for link l , $P^{avg}(l)$, is

$$P^{avg}(l) = \limsup_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m P_k(l).$$

Define the average network power vector as $\vec{P}^{avg} = (P^{avg}(1), P^{avg}(2), \dots, P^{avg}(L_E))$.

There may or may not exist a sequence of network power vectors $\vec{P}_1, \vec{P}_2, \dots$ that satisfy (6) and (7). If there does exist a sequence of such network power vectors, our aim is to minimize a linear function of \vec{P}^{avg} . An example of such a linear function is simply the total average power. We will consider the slightly more general cost function.

$$h(\vec{P}^{avg}) = \sum_{l \in \mathcal{E}} \alpha(l) P^{avg}(l), \quad (8)$$

where $\alpha(l)$ is a positive weight. For example, $h(\vec{P}^{avg})$ might equal the average power incident at a given point in space, or be equal to the total average power emanating from a closed boundary surrounding the network. Alternatively, $\alpha(l)$ might model the cost of a battery that drives transmitter $T(l)$.

A. Primal Problem

One of the main aims of this paper is to present an approach to *compute a schedule* of network power vectors $\vec{P}_1, \vec{P}_2, \dots$, satisfying the peak power constraints (6) and the minimum average rate constraints (7), such that $h(\vec{P}^{avg})$ is minimized. This primal problem is succinctly stated as

$$\min\{h(\vec{P}^{avg})\} \text{ subject to (6) and (7).} \quad (9)$$

Note this optimization involves choosing optimal power levels in each slot for each transmitter. Using a duality approach below, we reduce the problem to a convex optimization problem over a *single* slot.

B. Duality Approach

Let the value of the optimal cost in the problem (9) as a function of \vec{C} be denoted by $H(\vec{C})$. If for a given value of \vec{C} , no schedule of network power vectors exists satisfying (6) and (7), define $H(\vec{C}) = +\infty$.

Note that $H(\vec{C})$ is a convex function of the vector \vec{C} . To see this, suppose that $\vec{C} = \lambda\vec{C}_1 + (1 - \lambda)\vec{C}_2$, where $0 < \lambda < 1$. Let π_i be an optimal policy that supports the average data rates specified by \vec{C}_i , i.e. it achieves minimum cost $H(\vec{C}_i)$. Roughly speaking, we can create a new policy π by "time sharing" between the policies π_1 and π_2 , such that policy π_1 is used a fraction λ of the time and policy π_2 is used a fraction $1 - \lambda$ of the time. The policy π supports average data rate rates that are at least as large as \vec{C} . Furthermore, by the linearity of $h(\cdot)$, the cost of policy π is $\lambda H(\vec{C}_1) + (1 - \lambda)H(\vec{C}_2)$. By definition of $H(\vec{C})$, the cost of policy π is at least $H(\vec{C})$. Thus we have $H(\vec{C}) \leq \lambda H(\vec{C}_1) + (1 - \lambda)H(\vec{C}_2)$.

Define a set of dual variables $\vec{\beta} = \{\beta(1), \beta(2), \dots, \beta(L_E)\}$ for each link. Define a "potential" function V according to

$$V(\vec{P}, \vec{C}, \vec{X}) = h(\vec{P}) + \sum_{l \in \mathcal{E}} \beta(l) [C(l) - X(l)]$$

The dual objective function is defined as

$$g(\vec{\beta}) = \min\{V(\vec{P}^{avg}, \vec{C}, \vec{X}^{avg})\} \text{ subject to (6).}$$

Note the absence of the minimum rate constraints (7) in the definition of $g(\vec{\beta})$. For any non-negative vector $\vec{\beta}$, using (7), note that

$$\begin{aligned} H(\vec{C}) &\geq \min\{V(\vec{P}^{avg}, \vec{C}, \vec{X}^{avg})\}, \text{ subject to (6) and (7)} \\ &\geq \min\{V(\vec{P}^{avg}, \vec{C}, \vec{X}^{avg})\}, \text{ subject to (6)} \\ &= g(\vec{\beta}). \end{aligned} \quad (10)$$

Thus, for any non-negative vector $\vec{\beta}$, the minimum value of the objective function in the optimization problem (9) is lower bounded by $g(\vec{\beta})$. This observation leads to the dual optimization problem,

$$\max\{g(\vec{\beta}) : \vec{\beta} \geq \vec{0}\}. \quad (11)$$

A geometric interpretation of the inequality (10) can be obtained by considering the surface of $H(\cdot)$ in $(L_E + 1)$ -dimensional Euclidean space. In particular, $H(\vec{C})$ is lower bounded by considering a supporting hyperplane that is normal to the vector $(\vec{\beta}, 1)$. Since $H(\cdot)$ is convex, there exists a supporting hyperplane which passes through the point $(\vec{C}, H(\vec{C}))$. This is simply the classical argument used to prove that there is no "duality gap," [2] i.e.

$$\begin{aligned} H(\vec{C}) &= \max\{g(\vec{\beta}) : \vec{\beta} \geq \vec{0}\} \\ &= g(\vec{\beta}^*). \end{aligned} \quad (12)$$

Here, the supporting hyperplane passing through the point $(\vec{C}, H(\vec{C}))$ is normal to the vector $(\vec{\beta}^*, 1)$. From this geometric interpretation, it can be seen that the components of the optimal dual variable vector $\vec{\beta}^*$ represent a sensitivity of the optimal cost with respect to a perturbation in the minimum average data rate for a link, i.e.

$$\beta^*(l) = \frac{\partial H(\vec{C})}{\partial C(l)}. \quad (13)$$

Computation of $g(\vec{\beta})$ involves optimizing over all schedules of network power vectors satisfying the peak power constraint in every slot. However, since the potential function $V(\vec{P}, \vec{C}, \vec{X})$ is linear in \vec{P}, \vec{C} , and \vec{X} , it follows that $g(\vec{\beta})$ can be computed by an optimization over a single slot, i.e.

$$g(\vec{\beta}) = \min\{V(\vec{P}, \vec{C}, \vec{X}) : \vec{P} \in S_P\},$$

where $\vec{X} = \vec{X}(\vec{P})$ is defined as in (2) and (1), and S_P is defined as the polyhedral set $S_P = \{\vec{P} : \sum_{l \in \mathcal{E}(i)} P(l) \leq P^{max}(i) \text{ for all } i \text{ and } 0 \leq P(l) \text{ for all } l \in \mathcal{E}\}$. Let M be the number of extreme points of S_P , and let \vec{P}_m^{ext} , $m = 1, 2, \dots, M$ denote the extreme points of S_P . By definition, each point in S_P can be represented as a convex

combination of \vec{P}_m^{ext} , $m = 1, 2, \dots, M$. An upper bound on M is 2^{L_E} . If M_i is the number of links emanating from node i , then $M = \prod_{i=1}^N (1 + M_i)$.

Even though it can be verified that $V(\vec{P}, \vec{C}, \vec{X}(\vec{P}))$ is *not* a concave function of the vector \vec{P} , it is easy to show by taking derivatives that $V(\vec{P}, \vec{C}, \vec{X}(\vec{P}))$ is a concave function of $P(l)$ for each $l \in \mathcal{E}$. A standard convexity argument then shows that

$$g(\vec{\beta}) = \min\{V(\vec{P}_m^{ext}, \vec{C}, \vec{X}(\vec{P}_m^{ext})) : 1 \leq m \leq M\}.$$

Thus, the complexity of computing $g(\vec{\beta})$ is $O(M)$. In fact, note that $V(\vec{P}_m^{ext}, \vec{C}, \vec{X}(\vec{P}_m^{ext}))$ is an affine function of $\vec{\beta}$, which implies that the dual objective function $g(\vec{\beta})$ is the *pointwise minimum* of M affine functions, i.e.

$$g(\vec{\beta}) = \min_{1 \leq m \leq M} \{a'_m \vec{\beta} + b_m\}. \quad (14)$$

We have developed an iterative ascent algorithm to solve the dual problem (11). Exploiting the piecewise linearity of $g(\vec{\beta})$, our algorithm converges in a finite number of iterations. Alternately, by introducing a slack variable, the dual problem (11) can be reduced to a linear program and be efficiently solved using interior point methods.

If the required minimum rate vector \vec{C} is infeasible, i.e. $H(\vec{C}) = +\infty$, it can be seen that $g(\vec{\beta})$ is unbounded from above over the set of all $\vec{\beta} \geq 0$. If \vec{C} is feasible, it is easy to see that

$$H(\vec{C}) \leq \sum_{i=1}^N P^{max}(i).$$

Thus we can safely declare \vec{C} infeasible if we find a value of $\vec{\beta}$ such that $g(\vec{\beta}) > \sum_{i=1}^N P^{max}(i)$.

C. Computing the Optimal Policy

Solving the dual problem (11) yields the optimal dual variable vector $\vec{\beta}^*$ and extremal network power vectors $\mathbf{P}^{*,i}$, such that

$$\begin{aligned} g(\vec{\beta}^*) &= \min\{g(\vec{\beta}) : \vec{\beta} \geq 0\} \\ &= V(\vec{P}^{*,i}, \vec{C}, \vec{X}(\vec{P}^{*,i})). \end{aligned} \quad (15)$$

The vectors $\mathbf{P}^{*,i}$ are extreme points of S_P that represent optimal transmission modes and the power levels of the transmitters in those transmission modes. In other words, assuming $H(\vec{C})$ is finite, an optimal schedule of network power vectors exists that consists solely of these extremal network power vectors $\mathbf{P}^{*,i}$. This means that in the optimal policies we find, in each slot, every node is either not transmitting at all, or is transmitting at the maximum possible peak power to exactly one receiver.

Moreover, since a hyperplane in $(L_E + 1)$ dimensional Euclidean space is determined by $L_E + 1$ linearly independent points contained within it, an optimal policy can be constructed that consists of at most $L_E + 1$ extremal network power vectors $\mathbf{P}^{*,i}$. Let K be the number of extremal network power vectors $\mathbf{P}^{*,i}$ such that (15) holds. Define $\hat{K} = \min\{K, L_E + 1\}$. We thus have

$$H(\vec{C}) = g(\vec{\beta}^*) = V(\vec{P}^{*,i}, \vec{C}, \vec{X}(\vec{P}^{*,i})),$$

for all i such that $1 \leq i \leq \hat{K}$.

It is easy to see that in any optimal policy (assuming $H(\vec{C})$ is finite), the average rate on each link l , $X^{avg}(l)$, is exactly equal to $C(l)$. Otherwise, we could decrease the total average power further without violating any rate constraints thus contradicting the optimality of the solution. Let $X^{*,i}(l)$ denote the rate of link l corresponding the optimal extremal network power vector $\mathbf{P}^{*,i}$. Assuming $H(\vec{C})$ is finite, it can be shown that (12) guarantees the existence of a "weight vector" $\vec{\lambda}^* = \{\lambda^*(1), \lambda^*(2), \lambda^*(\hat{K})\}$ such that the equations (16) and (17) below are satisfied:

$$\sum_{i=1}^{\hat{K}} \lambda^*(i) X^{*,i}(l) = C(l) \text{ for all links } l \in \mathcal{E} \quad (16)$$

$$\sum_{i=1}^{\hat{K}} \lambda^*(i) = 1 \text{ where } \lambda^*(i) \geq 0 \quad (17)$$

Note that $\hat{K} \leq L_E + 1$, so the above system of equations is always overdetermined. There must be $\hat{K} + 1$ of these equations which are linearly independent. Once we have identified such a set of equations, we solve them by inverting a $(\hat{K} + 1) \times (\hat{K} + 1)$ matrix.

The value of $\lambda^*(i)$ indicates the relative frequency at which the extremal network power vector $\mathbf{P}^{*,i}$ is utilized in an optimal policy. This then determines a family of optimal policies, which can be constructed by alternating between the extremal network power vectors in accordance with the weight vector $\vec{\lambda}^*$. A specific optimal policy, for example, can be constructed using a pseudo-random number generator to select which extremal network power vector is used in each slot, such that the network power vector $\mathbf{P}^{*,i}$ is used with probability $\lambda^*(i)$. Alternatively, a specific optimal schedule can be constructed in such a way as to provide "smooth service" to each link, though this topic is beyond the scope of the present paper.

D. Reducing Complexity

The algorithm developed to solve the dual problem considers all M transmission modes. However, in the optimal schedule we have no more than $L_E + 1$ transmission

modes. Since M can be as large as 2^{L_ϵ} , the complexity of minimizing the dual objective function can be exponential. Thus, reducing the set of possible transmission modes considered can greatly reduce complexity. A natural consequence of our algorithm for a network of nodes with omni-directional antennas is that, no node transmits or receives data at the same time in the optimal schedule due to high self interference. This observation can reduce the number of candidate transmission modes significantly. In general, in many cases it is possible using heuristics to eliminate many "obviously inefficient" transmission nodes from consideration, since they are unlikely to be used in an optimal policy. This idea is similar in spirit to the approach used in [5] to construct a set of "valid" links. The resulting policies we find can be shown to be optimal among all policies which are constrained to use only the set of chosen candidate transmission modes.

III. ROUTING AND CAPACITY ALLOCATION

In the previous section, we presented an algorithm that finds an optimal link scheduling and power control policy to support a *given* traffic rate on each link in the network. In this section, we build on this to construct a policy which *allocates* traffic rates on each link in accordance with a routing algorithm, given a traffic matrix which describes the rate at which traffic is to be carried between specific source-destination pairs. The routing algorithm is guided by the computation of the minimum cost associated with optimal scheduling and power control for supporting traffic rates on each link that have already been allocated, as described in the previous section.

For example, consider a series of iterations, where at a given iteration we wish to support a small amount of additional traffic from one of the given source-destination pairs. Initially, no traffic is supported, i.e. the required minimum data rate on each link is zero. We compute the optimal values for the dual variables, β^* . Using (13), we can estimate the cost of supporting additional traffic on each link using a first order Taylor approximation. For example, suppose we wish evaluate a given route r . Let \vec{r} be the L_ϵ dimensional vector, such that the component $r(l) = 1$ if link l is included in the route \vec{r} , and $r(l) = 0$ otherwise. If we route an additional ϵ units of traffic along route r , the required data rate vector will increase from \vec{C} to $\vec{C} + \epsilon \vec{r}$. Using (13), the cost of supporting the additional ϵ units of traffic along route r can be determined from the approximation

$$H(\vec{C} + \epsilon \vec{r}) = H(\vec{C}) + \epsilon \sum_{l \in \mathcal{L}} \beta^*(l) r(l).$$

Therefore, in order to determine the best possible path along which an additional increment of traffic should be routed, we can use a shortest path algorithm (such as Dijkstra's algorithm or the Bellman-Ford algorithm), where we set the weight for link l to the sensitivity value $\beta^*(l)$. After we have identified the shortest path, we attempt to support the additional traffic along this path. If the new value of \vec{C} is infeasible, we either reduce the step-size ϵ until the additional traffic can be supported, try supporting additional traffic for another source-destination pair, or otherwise reach a prespecified termination condition.

In the examples, we will present below, it is apparent that $H(\cdot)$ is piecewise linear, which results in "sharp corners". When we use the above algorithm, it is possible that the link sensitivities can change abruptly. This can be partially addressed by choosing the step size ϵ sufficiently small. We can also "backtrack" by re-adjusting the rate allocations after each iteration, if necessary, to shift traffic towards shorter paths.

Indeed, since $H(\cdot)$ is convex, we can use the optimal routing framework developed by Bertsekas et al. (see [3]). In this class of algorithms, any feasible routing is first chosen. By incrementally adjusting the routing in an iterative manner, flows are shifted towards shortest paths at each iteration, and eventually a globally optimum routing is obtained. There are many possible variations, for example regarding whether flow is shifted simultaneously for every source destination pair, or whether flow is shifted one at a time for each source destination pair. In the present context, we are faced with a more difficult problem, since apparently $H(\cdot)$ is not differentiable everywhere. In addition, since evaluating the function $H(\cdot)$ is not trivial, it is more difficult to ascertain whether a given routing is feasible, i.e. the capacity of a given link depends on the rate carried by other links. Nevertheless, our initial numerical experiments that are based on (13) and framework developed by Bertsekas et. al. suggest that such an approach may be quite useful in practice.

As expected, since $H(\cdot)$ is convex, we find that $\beta^*(l)$ increases after we allocate additional traffic on link l . Interestingly, we find that sometimes $\beta^*(l)$ will *decrease* when we allocate additional traffic on a link $l' \neq l$. This corresponds to negative cross partial derivatives of $H(\cdot)$, and does not contradict the fact that $H(\cdot)$ is convex and non-decreasing. As a result of this, routes which initially appear unattractive sometimes become more attractive after traffic is added on other links. As we will see, this phenomenon can lead to traffic from the same source-destination pair being routed over multiple paths, which is somewhat counterintuitive since all links share a common frequency band.

We mark that the approach of using optimal values of the dual variables as link weights in shortest path routing has also been used before in the context of revenue maximization in wired networks, by Mitra et. al. [11]. This work focuses on the asymptotic behavior of the link sensitivities with respect to uniform increases in demand rates. The authors of this previous work also observed the optimality of route diversification for price elastic traffic associated with maximizing network revenue.

The potential energy efficiency of a given link can be evaluated, *assuming* that all other links are idle. For our model, the minimum transmission energy required per bit on a link (i, j) is inversely proportional to $G(i, j)$. A minimum energy path is defined as a path which has the smallest possible energy required to transport a fixed amount of data, *irregardless* of the delay associated with transporting that data. Previous work on routing in wireless networks (for example see [14][13]) is focused on energy efficiency without regard to delay, i.e. such work proposes using minimum energy paths.

In our framework, we find that the paths over which traffic is allocated in some optimum routings does *not* always correspond to using minimum energy paths. However, when required data rates are sufficiently low, our routing framework will apparently always use minimum energy paths. This is illustrated by example in the next section.

IV. RESOURCE ALLOCATION EXAMPLES

In this section, we illustrate the gains of using our wireless resource allocation framework through a series of network topologies. We focus on simple examples in an attempt to gain insight from optimal allocations.

A. Example 1: String Topology

A string topology consists of a row of nodes connected by means of directed links as shown in Figure 1. This topology represents a “transport” network carrying data over long distances. Node 1 is the source of data and node 5 is the sink. Each node in this example has an omnidirectional antenna, and all nodes have identical peak transmission power constraints of 1 Watt. The path-loss between nodes was modeled as an inverse square of the distance for all the examples in this paper, i.e. $G(i, j) = \frac{1}{d(i, j)^2}$. The ambient noise power at all the nodes is assumed to be constant. We compare the maximum throughput of the single session between node 1 to node 5 achieved by our link scheduling policy to a TDMA scheduling policy. The TDMA policy schedules exactly one link at a time, activating each link at maximum power for a *fourth* of the

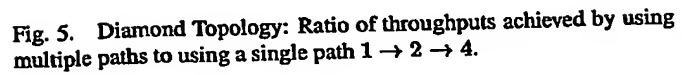
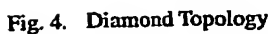
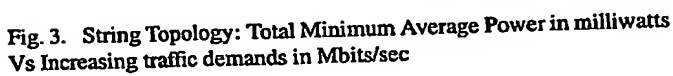
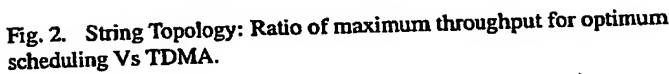
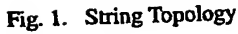
time. We plot the ratio of throughputs achieved by the our link scheduling algorithm to TDMA with increasing values of ambient noise in Figure 2. The ambient noise is normalized with respect to the peak received power (from a single transmitter) in the network and is scaled by $\log_2(\cdot)$ (for all graphs in this section). The maximum throughput achieved by our policy is higher than TDMA for all values of ambient noise. In low noise regimes, TDMA is a near-optimal policy. However, as the level of ambient noise at nodes increases, the gains in throughput of our policy increase significantly. The optimal policy schedules multiple transmissions at the same time, even in fairly low ambient noise regimes, thereby exploiting spatial reuse and thus outperforming TDMA. This effect is more pronounced in moderate and high ambient noise regimes where the optimal policy schedules concurrent transmissions even though they are in close geographic proximity.

To get a sense of how energy efficient the optimal policy is as it supports increasing traffic demands, we plot the total average power with increasing traffic rates for the session between node 1 and node 5 in Figure 3. For this experiment, we fix the level of the logarithm of the normalized value of ambient noise to -0.67 . This value corresponds to the 3rd data point on the x-axis in Figure 2. We also set $W' = 10^7$.

Figure 3 indicates that the total minimum average power is a linear function of the traffic demand for average rates below 4 Mbits/sec. Below 4 Mbits/sec, our scheduling policy reduces to TDMA, with only one link transmitter active at any time. As the traffic demand of the session increases beyond 4 Mbits/sec, the total average power increases at a faster rate. For traffic load between 4 Mbits/sec and 4.98 Mbits/sec, the transmission modes in the optimal policy are given by $\{(1, 2)\}$, $\{(2, 3)\}$, $\{(3, 4)\}$ and $\{(1, 2), (4, 5)\}$. For traffic loads above 4.98 Mbits/sec to the maximum throughput possible, the optimal transmission modes are $\{(1, 2)\}$, $\{(2, 3)\}$, $\{(3, 4)\}$, $\{(1, 2), (3, 4)\}$ and $\{(1, 2), (4, 5)\}$. We see that in order to support high traffic loads, the optimal link scheduling policy activates a large number of links simultaneously. Note that the transmission modes in the optimal policy are half-duplex. Therefore, transmissions in the optimal schedule cannot consume a total average power of more than 2000 mW for this topology.

B. Example 2: Diamond Topology

In this experiment, we consider an asymmetric diamond topology as shown in Figure 4 to illustrate the sub-



We compare the ratio of throughputs for these policies in Figure 5. Clearly, using multiple paths yields higher throughputs for all values of ambient noise. The increase in throughput by splitting traffic over multiple paths is significant even for moderate levels of ambient noise. This result is somewhat surprising, since all links are sharing a common bandwidth and we have a per-node peak power constraint. An intuitive explanation of this is that by splitting the traffic over both paths, the transmission modes $\{(1, 2), (3, 4)\}$ and $\{(2, 4), (1, 3)\}$, which do not have much interference, can be alternated in time. This results in substantial delivery of data to the destination in every slot. If we only allow the path $1 \rightarrow 2 \rightarrow 4$, in order to avoid self interference, the links $(1, 2)$ and $(2, 4)$ must be active one at a time, so that node 4 will only be delivered data in every other slot.

To illustrate energy efficiency issues, we plot the total minimum average power with increasing traffic demand at node 1 in Figure 6. The logarithm of the normalized value

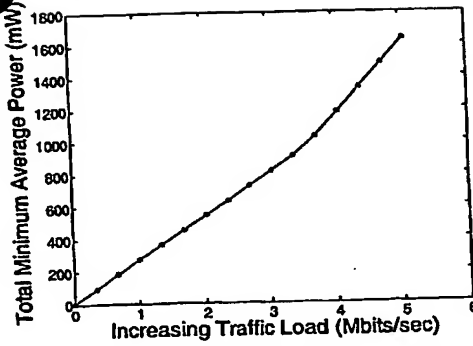


Fig. 6. Diamond Topology: Total Minimum Average Power in milliwatts Vs Increasing traffic demands in Mbits/sec

of ambient noise for this point is 0.45 and corresponds to the 3rd data point on the x-axis in Figure 5. We set $W' = 10^7$. The maximum data rate achieved by TDMA using the minimum energy path $1 \rightarrow 2 \rightarrow 4$ is 3.65 Mbit/sec. In comparison, our policy is capable of supporting a data rate of up to 5.05 Mbits/sec by using both the minimum energy path as well as the non-minimum energy path. For traffic loads below 3.65 Mbits/sec, our policy is identical to TDMA over the path $1 \rightarrow 2 \rightarrow 4$. As the traffic load increases beyond 3.65 Mbits/sec, the optimal policy uses starts using both paths by scheduling transmission modes $\{(1, 2), (3, 4)\}$ and $\{(1, 3), (2, 4)\}$ in addition to TDMA transmission modes. The average power required to achieve higher traffic demands beyond 3.65 Mbits/sec increases with a greater slope as can be seen in Figure 6. The maximum throughput of 5.05 Mbits/sec is achieved by splitting traffic by using both paths and scheduling transmission modes $\{(1, 2)\}$, $\{(2, 4)\}$, $\{(1, 2), (3, 4)\}$ and $\{(1, 3), (2, 4)\}$.

C. Example 3: Hierarchical Topology

A hierarchical topology as shown in Figure 7 represents a complex topology such as an access network, where the number of nodes in each tier decrease as we get closer to the access point node 6. This experiment highlights the significant gains in throughput by jointly routing and scheduling to an alternative policy that optimally schedules links over pre-determined *minimum energy routes*. Nodes 1, 2 and 3 are sources of data and node 6 is the sink. We assume $G(1, 6) = \frac{1}{d(1,6)^4}$ and $G(2, 6) = \frac{1}{d(2,6)^4}$; the path loss between all other nodes is given by the inverse square law of distance. The minimum energy paths for the source nodes are: $\{1 \rightarrow 3 \rightarrow 6\}$, $\{2 \rightarrow 3 \rightarrow 6\}$ and $\{3 \rightarrow 6\}$ respectively. Our integrated routing and scheduling algorithm is allowed to utilize all possible paths to route their data to node 6. We constrain the rates

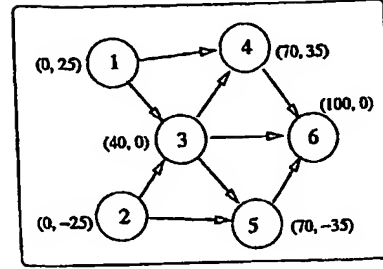


Fig. 7. Hierarchical Topology: Access Network

of each source to be identical. We plot the ratio of the total maximum throughput of our routing policy to that achieved by using only the minimum energy paths, as a function of ambient noise in Figure 8.

We find that at low noise levels, our policy supports a throughput at least 30% more than the data rate achieved using minimum energy paths alone. As the level of ambient noise increases, our policy supports a significantly higher throughput. The base policy schedules links in a TDMA sequence for all values of ambient noise. The optimal policy schedules concurrent links even for low levels of ambient noise. In the optimal policy, nodes 1 and 2 route their traffic over non-minimum energy paths $1 \rightarrow 4 \rightarrow 6$ and $2 \rightarrow 5 \rightarrow 6$ respectively, apart from the minimum energy paths, for all values of ambient noise. For low ambient noise, the optimal policy schedules the transmission modes $\{(1, 3)\}$, $\{(2, 3)\}$ and $\{(3, 6)\}$ according to TDMA but also allows modes $\{(1, 4), (5, 6)\}$ and $\{(2, 5), (4, 6)\}$ to be active for reasonable fractions of time. As the level of ambient noise increases, the optimal policy splits node 3's traffic over paths $3 \rightarrow 4 \rightarrow 6$ and $3 \rightarrow 5 \rightarrow 6$. In fact, for sufficiently high ambient noise, the optimal policy schedules a large number of concurrent links, the dominant transmission modes being $\{(1, 4), (5, 6), (2, 3)\}$ and $\{(2, 5), (4, 6), (1, 3)\}$ and $\{(1, 4), (2, 5), (3, 6)\}$.

V. CONCLUSION

We have developed an integrated routing, link scheduling and power allocation policy for a general multihop network that minimizes the total average power consumption to support minimum average rate requirements per link. Our policy can support higher throughputs than with conventional approaches to radio resource allocation, at the expense of decreased energy efficiency. Our policy requires time synchronization between transmitters, and requires that channel conditions remain constant over several time slots. It appears to be most appropriate in the

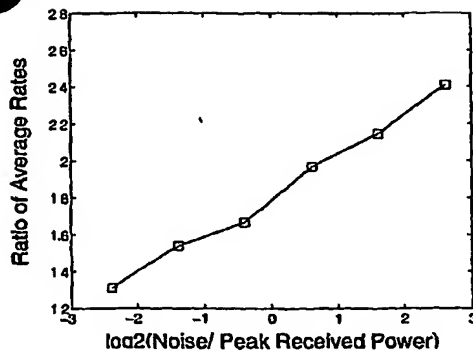


Fig. 8. Hierarchical Topology: Ratio of maximum throughputs achieved by using multiple paths to that using minimum energy paths only

context of providing a means for wireless interconnection of fixed stationary nodes, such as wireless access nodes.

We find that our optimal power control policies are such that each node is either transmitting at peak power to a single receiver or not transmitting at all, which makes it simpler to implement. This also means that nodes can go to "sleep" much of the time, when they are not transmitting or receiving, which potentially conserves a great deal of energy [4]. The optimal link schedule time-shares a small number of optimal subsets of links ($\leq L_E + 1$) in order to achieve the required data rates. The collection of optimal transmission modes depend on the level of ambient noise and the capacity allocations on links. The optimal policy for low required data rates or in low ambient noise regimes schedules links in a TDMA sequence. In moderate or high noise regimes, or to achieve higher data rates, the optimal strategy involves scheduling multiple simultaneous transmissions even though they may be in close geographic proximity.

A byproduct of solving the scheduling and power control problem yields link sensitivities that tell us the change in the minimal total average power with respect to a perturbation of the required data rate on a link. The minimal required average power is a convex function of the required data rates on each link. This leads us to a routing algorithm that is based on previously developed non-linear programming techniques.

We find that using minimum-energy paths to route traffic and TDMA to schedule links is *near-optimal* in low noise regimes and in the case where the required average data rates are small. As the level of ambient noise increases, or the required average data rates increase, our optimal routing policy exploits the use of multiple paths to support higher data rates, and the optimal scheduling policy makes use of concurrent transmissions on links. We find that non-minimum energy paths can be exploited to

increase throughput, despite the fact that all links share a common bandwidth.

We remark that it is easy to generalize our problem formulation in a number of directions. For example, we can specify a maximum peak power constraint for an arbitrary collection of nodes, rather than per each node. This may be useful to limit signal emissions from an entire network or subnetwork.

One topic we are currently investigating is the use of hierarchical scheduling, whereby links are partitioned into clusters that are geographically remote. The advantage of this is that the clusters can be scheduled independently, with only a loose coupling between different clusters. The signal emissions from a given cluster can be modeled as "ambient noise" for purposes of scheduling transmissions in another remote cluster.

REFERENCES

- [1] N. Bambos, S. Chen and G. Pottie, "Radio Link Admission Algorithm for Wireless Networks with Power Control and Active Link Quality Protection", *Proc. IEEE Infocom'95*, pp. 97-104, 1995.
- [2] D. Bertsekas, "Non-linear Programming", *Athena Scientific*, 1999.
- [3] D. Bertsekas and J. Tsitsiklis, "Parallel and Distributed Computation: Numerical Methods", *Prentice Hall* 1989.
- [4] B. Chen, K. Jamieson, R. Morrison and H. Balakrishnan, "Span: An Energy-Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks", *Proc. ACM MOBICOM Conference*, July 2001.
- [5] T. Elbatt, A. Ephremides, "Joint Scheduling and Power Control for Wireless Ad-hoc Networks", *Proc. IEEE Infocom 2002*.
- [6] A. Ephremides, T. V. Truong, "Scheduling Broadcasts in multi-hop radio networks", *IEEE Trans. on Communications* v.38, Apr. 1990.
- [7] G. Foschini, Z. Miljanic, "A Simple Distributed Autonomous Power Control Algorithm and its Convergence", *IEEE Transactions on Vehicular Technology*, vol. 42 no. 4, pp 641-646, Nov. 1993.
- [8] P. Gupta, P. R. Kumar, "The Capacity of Wireless Networks", *IEEE Transactions on Information Theory*, vol. IT-46, no. 2, pp. 388-404, March 2000.
- [9] L. Kleinrock and J. A. Silvester, "Optimum Transmission Radii in Packet Radio Networks or Why Six is a Magic Number", *National Telecommunications Conference*, Birmingham, Alabama, pp. 04.3.1-5, December 1978.
- [10] H. Luo, S. Lu, V. Bharghavan, "A New Model for Packet Scheduling in Multihop Wireless Networks", *ACM Mobicom*, August 2000.
- [11] D. Mitra, K.G. Ramakrishnan and Q. Wang, "Combined Economic Modeling and Traffic Engineering: Joint Optimization of Pricing and Routing in Multi-Service Networks", *Proc. 17th International Teletraffic Congress* 2001.
- [12] T. S. Rappaport, "Wireless Communications: Principles and Practice", *Prentice Hall*.
- [13] V. Rodoplu and Teresa Meng, "Minimum energy mobile wireless networks", *Proc. ICC'98* vol. 3, pp. 1633-1639, June 1998.

- [14] S. Singh, M. Woo and C. S. Raghavendra, "Power-aware routing in mobile ad hoc networks", *Proc. of Fourth Annual ACM/IEEE Conference on Mobile Computing and Networking* pp. 181-190, Oct. 1998.
- [15] S. Toumpis, A. J. Goldsmith, "Capacity Regions for Wireless Adhoc Networks", *International Symposium on Communication Theory and Applications*, April 2001.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.